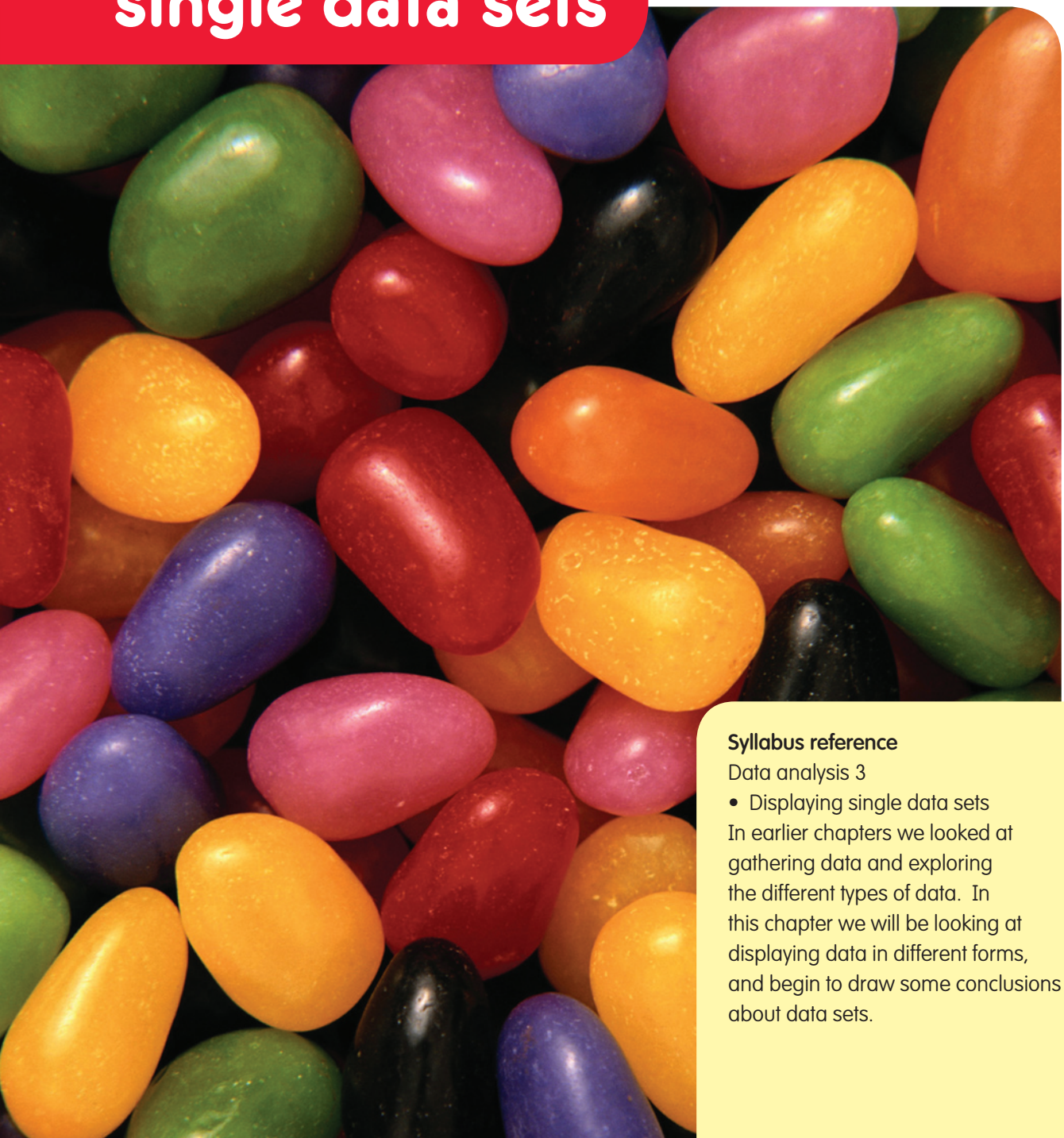
9

Displaying single data sets



Syllabus reference

Data analysis 3

- Displaying single data sets

In earlier chapters we looked at gathering data and exploring the different types of data. In this chapter we will be looking at displaying data in different forms, and begin to draw some conclusions about data sets.

ARE YOU READY?

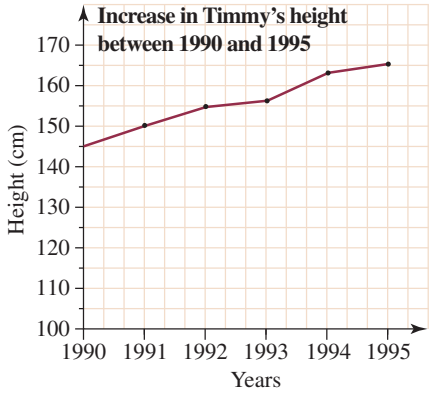
Try the questions below. If you have difficulty with any of them, extra help can be obtained by completing the matching SkillSHEET. Either click on the SkillSHEET icon next to the question on the *Maths Quest Preliminary Course* eBookPLUS or ask your teacher for a copy.

eBookplus
Digital doc
SkillSHEET 9.3
doc-1561
Reading a line graph

Reading a line graph

1 The line graph at right shows the height of a child (Timmy) over 5 years.

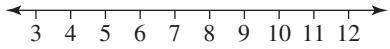
a How tall was Timmy at the start of the measurement period?
b How much did Timmy grow in the first year?
c How much did Timmy grow over the five years?
d How many years did it take for Timmy to grow 10 cm?



eBookplus
Digital doc
SkillSHEET 9.4
doc-1562
Presenting data as a dot plot

Presenting data as a dot plot

2 Copy and complete the dot plot at right for the data shown below.
6, 4, 3, 11, 9, 12, 5, 7, 7, 6, 3, 6, 5, 6, 8, 6, 9, 5, 6, 7

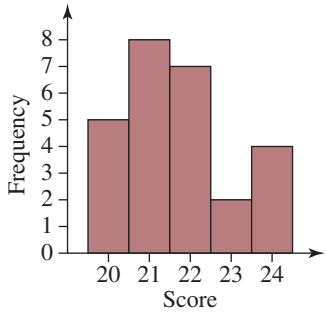


eBookplus
Digital doc
SkillSHEET 9.5
doc-1567
Producing a frequency table from a frequency histogram

Producing a frequency table from a frequency histogram

3 Copy and complete the following frequency table to show the data represented in the frequency histogram.

Score (x)	Frequency (f)
20	5
21	
22	
23	
24	



eBookplus
Digital doc
SkillSHEET 9.6
doc-1571
Finding the median of a small data set

Finding the median of a small data set

4 Find the median of each of the following data sets.

a 2, 6, 8, 4, 5, 6, 7
b 13, 10, 15, 12, 18, 17, 12, 12
c 52, 45, 23, 83, 9, 45, 71, 84, 90, 183
d 9.2, 9.3, 9.4, 9.3, 9.4, 9.5, 9.9, 9.4, 9.7, 9.8, 9.0, 10.0, 9.4, 9.2, 9.9

eBookplus
Digital doc
SkillSHEET 9.7
doc-1576
Presenting data as a stem-and-leaf plot

Presenting data as a stem-and-leaf plot

5 Copy and complete the following stem-and-leaf plot for the given data.
52, 34, 38, 45, 10, 33, 27, 19, 33, 46

Key: 2 | 3 = 23

Stem	*Leaf*
1	_ 9
2	_
3	3 _ 4 _
4	5 _
5	_

9A Frequency tables

A **frequency table** is used to tabulate statistical data. A well-designed table allows the person doing the statistical investigation to easily tally up the data. In this chapter we are concerned with the tallying of quantitative data.

When working with quantitative data, each piece of data is known as a **score**. Quantitative data may be presented as either **grouped** or **ungrouped** data. The table at right shows the number of cinema visits during a month by each of 20 students. It is an example of ungrouped data because each score is a separate group in itself.

No. of visits	Frequency
0	6
1	7
2	4
3	2
4	1

Ungrouped data are suitable for discrete data that do not have a wide range of scores.

A frequency table usually has a tally column that allows us to enter a tally mark for each score as it is entered into the table. The gatepost method of tallying (the fifth tally mark crosses the previous four like a gatepost) allows us to easily count up the frequency at the end.

WORKED EXAMPLE 1

Fifty people were surveyed and asked the number of videos that they had hired from a video store in the past month. The results are shown below.

2	4	5	3	3	3	3	0	5	1
1	3	3	0	5	2	2	1	3	1
3	4	4	2	0	1	4	0	3	2
0	1	5	5	1	2	3	5	3	4
0	2	0	4	4	4	2	2	4	5

Enter the information in a frequency table.

THINK

1. Draw a table with three columns and with scores from 0 to 5.
2. Enter a tally mark for each score.
3. Count the tally marks for each score and enter the result in the frequency column.

WRITE

Score	Tally	Frequency
0	𝍸 II	7
1	𝍸 II	7
2	𝍸 IIII	9
3	𝍸 𝍸 I	11
4	𝍸 IIII	9
5	𝍸 II	7

When the data are continuous or if the data are spread over a wide range, it is useful to group the scores into *groups* or *classes*. The following table contains grouped data. It shows the number of passengers on each of 20 bus trips.

No. of passengers	Frequency
5–9	1
10–14	6
15–19	8
20–24	4
25–29	1

When making the decision to summarise some raw data by grouping them in a frequency table, the choice of class size (group size) is important. As a general rule, try to choose a class size so that 5 to 10 groups are formed. Choosing a suitable data label is also important.

Consider the following data that give the number of nails in each of a sample of 40 boxes.

130	122	118	139	126	128	119	124	122	123
132	138	129	139	116	123	126	128	131	142
137	134	126	129	127	118	130	132	134	132
137	124	134	134	120	137	141	118	125	129

Let's say that we wish to represent the data in a frequency table. It would be pointless trying to tally the data as individual (ungrouped) scores, as most scores would occur only once or twice in the whole distribution. It makes good sense to group the scores. A class size of five would be appropriate as this leads to the formation of six classes.

With grouped data, an extra column is shown for the class centre. The class centre is the middle of each class.

No. of nails	Class centre	Tally	Frequency
115–119	117	𝍸	5
120–124	122	𝍸 ||	7
125–129	127	𝍸 𝍸	10
130–134	132	𝍸 𝍸	10
135–139	137	𝍸 |	6
140–144	142	||	2

WORKED EXAMPLE 2

The height of 40 students was measured and the results are shown below.

146	**141**	**155**	**166**	**168**	**158**	**169**	**164**	**141**	**154**
159	**143**	**152**	**156**	**146**	**146**	**161**	**150**	**141**	**153**
152	**148**	**152**	**142**	**162**	**159**	**141**	**151**	**169**	**169**
164	**148**	**168**	**169**	**146**	**162**	**151**	**150**	**143**	**140**

Put the above results into a frequency table.

THINK

1. The data range from 140 cm to 168 cm. Choose a group of 5 cm beginning at 140 cm.
2. Calculate the class centres.
3. Draw a frequency table with four columns and room for six classes.
4. Enter a tally mark in the appropriate class as each height is read.
5. Complete the frequency table by counting the tally marks.

WRITE

Height	Class centre	Tally	Frequency
140–144	142	𝍸 III	8
145–149	147	𝍸 I	6
150–154	152	𝍸 IIII	9
155–159	157	𝍸	5
160–164	162	𝍸	5
165–169	167	𝍸 II	7

In some examples the class groupings overlap at the boundary. For example, if we are collecting information on people's mass, the categories used may be 50–55, 55–60, 60–65, etc. In such examples, the class centres would be 52.5, 57.5, 62.5, . . . etc.

When tabulating such data, convention is that a score on the boundary is placed in the higher class. For example, a person with a mass of 60 kg would be placed in the 60–65 class.

REMEMBER

1. Statistical information may be tabulated using a frequency table.
2. A frequency table has three columns for score, tally and frequency.
3. The data can be entered as either grouped or ungrouped data.
4. Ungrouped data are where each score becomes a group on its own and is suitable for a small range of scores for which the data are discrete.
5. Grouped data are put into classes. They are suitable for continuous data or discrete data for which there is a wide range of scores. For grouped data the class centre should be shown.

EXERCISE 9A

Frequency tables

eBookplus
Digital doc
SkillSHEET 9.1
doc-1558
Presenting data in a frequency table

eBookplus
Digital doc
EXCEL Spreadsheet
doc-1559
Frequency tables

1 WE1 Twenty households were surveyed to find the number of people in that household. The results are shown below.

4 3 4 6 3 2 5 2 7 4 5 6 4 3 5 4 6 2 3 4

Use this information to complete a copy of the frequency table below.

Score	Tally	Frequency
2		
3		
4		
5		
6		
7		

2 The marks of 25 students on a spelling test are shown below.

4 5 8 5 10 7 6 9 7 6 5 7 6
4 7 5 6 7 5 7 8 4 6 8 7

Put this information into a frequency table.

3 The scores of 50 professional golfers in a round of golf are shown below.

72 70 69 75 78 68 66 68 67 72 72 71 68 73 72 71 74
77 72 73 72 72 72 74 70 71 73 72 77 74 76 68 69 68
77 75 72 72 72 71 73 72 70 74 72 71 73 68 67 67

Display this information in a frequency table.

4 WE2 A class of 30 students sat for a Mathematics test. Their results out of 100 are shown below.

68 72 58 45 69 92 38 51 70 65 69 73 52 76 48
69 73 41 42 73 80 50 60 49 65 94 88 85 53 60

Use these results to copy and complete the frequency table below.

Score	Class centre	Tally	Frequency
30–39			
40–49			
50–59			
60–69			
70–79			
80–89			
90–99			

5 A farmer measures the heights of his tomato plants. The results, in metres, are shown below.

0.93 1.21 2.03 1.40 1.17 1.53
1.82 1.77 1.65 0.63 1.24 1.99
0.80 2.14 1.53 2.07 1.96 1.05
0.94 1.23 1.72 1.34 0.75 1.17
1.50 1.41 1.74 1.86 1.55 1.42
1.52 1.39 1.76 1.67 1.28 1.43
2.13

Use the class groupings 0.6–0.8, 0.8–1.0, 1.0–1.2, . . . etc. to complete a frequency distribution table for these data.

6 The following data give the times (in seconds) taken for athletes to complete a 100 m sprint.

12.2 12.0 11.9 12.0 12.6 11.7
11.4 11.0 10.9 11.7 11.2 11.8
12.2 12.0 12.7 12.9 11.3 11.2
12.8 12.4 11.7 10.8 13.3 11.7
11.6 11.7 12.2 12.7 13.0 12.2

Construct a frequency distribution table for the data. Use a class size of 0.5 seconds.

Further development

7 Forty households were surveyed to find the number of people in each of them. The data was recorded as follows.

0,	3,	2,	4,	1,	2,	3,	2,	2,	2,
2,	1,	3,	4,	5,	2,	3,	1,	1,	1
0,	0,	2,	3,	4,	1,	3,	4,	2,	2,
0,	1,	2,	3,	2,	0,	2,	4,	5,	1

a Is the data discrete or continuous?

b Display the results in a frequency table.

8 A survey is conducted where 40 people were asked the number of hours of television that they watched each week.

a Is the data discrete or continuous?

b The results are as shown below.

10,	13,	7,	12,	16,	11,	6,	14,	6,	11,
5,	14,	12,	8,	27,	17,	13,	8,	14,	10
13,	7,	15,	10,	16,	8,	18,	14,	21,	28,
9,	12,	11,	13,	9,	13,	29,	5,	24,	11

Why is it more practical in this case to use class groupings rather than individual scores?

c Complete the frequency table below.

Class	Tally	Class centre	Frequency
5–10			
10–15			
15–20			
20–25			
25–30			

9 The number of phone calls made, on average, per week in a sample of 56 people is listed below.

21,	50,	8,	64,	33,	58,	35,	61,	3,	51,	5,	62,	16,	44,
56,	17,	59,	23,	34,	57,	49,	2,	24,	50,	27,	33,	55,	7,
52,	17,	54,	78,	69,	53,	2,	42,	52,	28,	67,	25,	48,	63,
12,	72,	36,	66,	15,	28,	67,	13,	23,	10,	72,	72,	89,	80

a Is the data discrete or continuous?

b Explain why classes of 1–10, 10–20, . . . etc. are not appropriate in this case.

c Complete a frequency table using a class size of 10.

10 Explain why continuous data should always be tabulated in class groupings.

11 A data set is tabulated in class groups of 10. If the data is discrete the class groupings would be 0–9, 10–19 etc., but if the data was continuous the class groupings would be 0–10, 10–20, . . . etc. What difference does this choice of class groupings make to the class centre?

9B Types of graphs

Once statistical information has been tabulated, the next step is to display the data in some type of graph. The type of graph that is chosen depends on the purpose of the graph.

Dot plots

A **dot plot** is used to display a set of scores on a number line. This graph is useful for showing a small number of scores.

WORKED EXAMPLE 3

Below are the scores out of 10 achieved by 11 students on a Maths quiz.

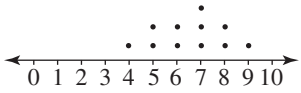
7 4 8 7 6 7 6 5 8 9 5

Show this information on a dot plot.

THINK

1. Draw a number line showing all numbers from 0 to 10.
2. Place a dot on the appropriate number for each score, stacking the dots to show where more than one of the same score has occurred.

WRITE



Sector graph (pie chart)

A **sector graph** is used to compare quantities. A circle is divided into proportions to be shown on the graph. To calculate the angles at the centre of the circle, the fraction of the total quantity is multiplied by 360°.

WORKED EXAMPLE 4

eBook*plus*
Tutorial
int-2323
Worked example 4

A poll was conducted on an issue of justice based on the statement 'Penalties for serious crime should be increased'. The responses were:

Agree 35 Disagree 20 Unsure 5.

Show this information in a sector graph.

THINK

1. Calculate the total number of responses.
2. Calculate the angles at the centre of the circle by multiplying each fraction by 360°.
3. Draw the graph.

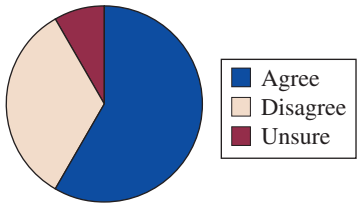
WRITE/DRAW

$35 + 20 + 5 = 60$

$\text{Agree} = \frac{35}{60} \times 360°$
$= 210°$

$\text{Disagree} = \frac{20}{60} \times 360°$
$= 120°$

$\text{Unsure} = \frac{5}{60} \times 360°$
$= 30°$



Line graph

A **line graph** is used to compare the change in one quantity with the change in another. It is suitable for quantitative data.

WORKED EXAMPLE 5

The table below shows the temperature taken in a school playground over a day.

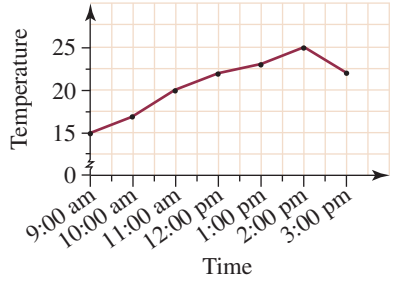
Time	9:00 am	10:00 am	11:00 am	12:00 pm	1:00 pm	2:00 pm	3:00 pm
Temp (°C)	15	17	20	22	23	25	22

Show this information in the form of a line graph.

THINK

1. Show time on the horizontal axis and temperature on the vertical axis.
2. Plot the points for the hourly temperature.
3. Join each point with a straight line.

DRAW



Bar and column graphs

A **bar graph** or a **column graph** is used when we want to show a quantity and compare it to other quantities. This is particularly suitable for categorical data.

In a bar graph, the categories are shown on the vertical axis and the results are shown on the horizontal axis. The bars are then drawn horizontally. These axes are reversed for a column graph.

WORKED EXAMPLE 6

The information shown below represents the unemployment rates in various NSW regions.

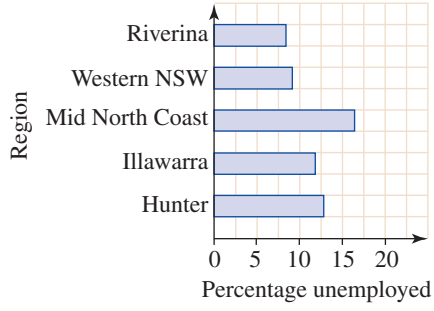
Hunter	**12.9%**
Illawarra	**11.8%**
Mid North Coast	**16.4%**
Western NSW	**9.1%**
Riverina	**8.4%**

Show this information in a bar graph.

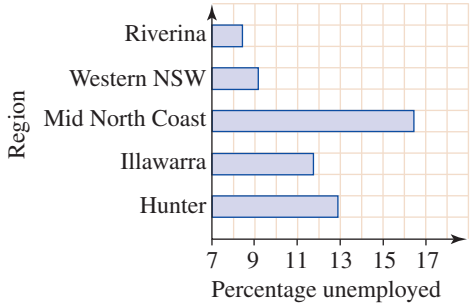
THINK

1. Draw the categories on the vertical axis and the percentages on the horizontal axis.
2. Draw horizontal bars for each category.

DRAW



Graphs can be used to misrepresent information. This can be done by adjusting the scale on the horizontal axis. The graph at right shows the same information as worked example 6.



Radar charts

A **radar chart** is similar to a line graph except it is circular. It is particularly suitable for showing data trends that repeat.

The 'radar' is drawn with the data being measured placed in equal sectors around the circle and the results having a scale emanating from the centre. The points are then plotted and joined.

WORKED EXAMPLE 7

eBookplus
Tutorial
int-2324
Worked example 7

The information below shows the sales in a department store over a year.

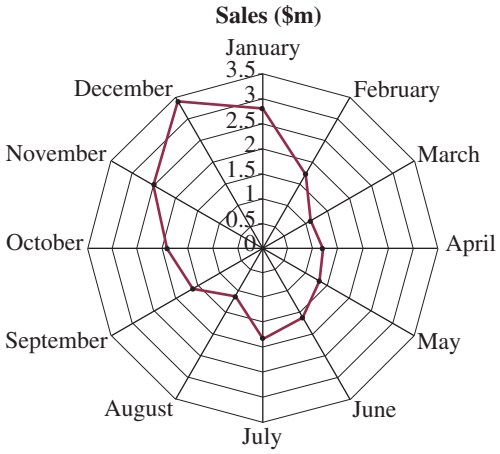
Month	Sales ($m)	Month	Sales ($m)
January	2.8	July	1.8
February	1.7	August	1.1
March	1.1	September	1.6
April	1.2	October	1.9
May	1.3	November	2.5
June	1.6	December	3.4

Show these data in a radar chart.

THINK

1. Draw equal sectors of 30°.
2. Draw the sales from the centre using 1 cm = $0.5 million.
3. Plot the points.
4. Join each point with a straight line.

DRAW



REMEMBER

1. A graph is used to visually display data that have been tabulated.
2. You should choose the most appropriate graph for the purpose of the display and the data type.
3. You should be able to draw dot plots, sector graphs, line graphs, bar graphs and radar charts.

EXERCISE 9B

Types of graphs

eBookplus
Digital doc
SkillSHEET 9.2
doc-1560
Reading a column graph

1 WE3 Below are the scores out of 10 on a spelling test as achieved by a small class of students. Show the information on a dot plot.

4 5 3 7 8 10 9 6 7 7 8 6 8 6 7

2 Below is the maximum temperature in Sydney each day during February.

28 35 33 34 30 27 28 29 30 26 28 30 31 33
32 31 30 31 29 29 28 25 26 30 31 30 29 29

Show this information on a dot plot.

eBookplus
Digital doc
SkillSHEET 9.3
doc-1561
Reading a line graph

3 WE4 During a season, a soccer team has 10 wins, 4 draws and 6 losses. Show these results in a sector graph.

4 Gary earns $600 per week. His weekly budget is:

Rent $200 Bills $150 Groceries $120 Car running costs $70.

The balance is put away for savings. Display Gary's budget in a sector graph.

eBookplus
Digital doc
SkillSHEET 9.4
doc-1562
Presenting data as a dot plot

5 WE5 The population over 8 years in a small town is shown in the table below.

Year	2000	2001	2002	2003	2004	2005	2006	2007
Pop.	12 000	12 500	13 250	13 500	13 600	14 100	15 000	16 000

Show this information in a line graph.

eBookplus
Digital doc
EXCEL Spreadsheet
doc-1563
Column graphs

6 Tania has a savings account. The balance of the account at the beginning of each month for a year is shown in the table below.

Month	Balance	Month	Balance
January	$200	July	$500
February	$270	August	$460
March	$320	September	$570
April	$260	October	$625
May	$380	November	$570
June	$430	December	$700

Show this information in a line graph.

eBookplus
Digital doc
EXCEL Spreadsheet
doc-1564
Line graphs

7 WE6 Below are the average weekly earnings of people in different categories of employment.

Labourers	$543	Plant operators	$598
Salespersons	$607	Clerks	$620
Tradespersons	$640	Para-professionals	$780
Professionals	$865	Managers	$906

Show this information in a horizontal bar graph.

eBookplus
Digital doc
EXCEL Spreadsheet
doc-1565
Bar graphs

eBook*plus*

Digital doc
EXCEL Spreadsheet doc-1566
Sector graphs

8 The table below shows the number of Olympic gold medals Australia has won since 1952.

Year and venue	No. of gold medals	Year and venue	No. of gold medals
1952 Helsinki	6	1980 Moscow	2
1956 Melbourne	13	1984 Los Angeles	4
1960 Rome	8	1988 Seoul	3
1964 Tokyo	6	1992 Barcelona	7
1968 Mexico City	5	1996 Atlanta	9
1972 Munich	8	2000 Sydney	16
1976 Montreal	0	2004 Athens	17

Show this information in a column graph.

9 WE7 The table below shows the average monthly temperature in Sydney.

Month	Temp (°C)	Month	Temp (°C)
January	28	July	15
February	30	August	16
March	26	September	20
April	24	October	22
May	20	November	25
June	17	December	27

Show this information in a radar chart.

10 The table below shows the percentage of televisions that are being watched over a 24-hour period.

Time	Percentage	Time	Percentage
12:00 am	12%	12:00 pm	30%
2:00 am	1%	2:00 pm	33%
4:00 am	2%	4:00 pm	45%
6:00 am	8%	6:00 pm	60%
8:00 am	15%	8:00 pm	78%
10:00 am	24%	10:00 pm	55%

Show this information in a radar chart.

Further development

11 **a** Describe the advantages of displaying data in a dot plot.

b The data below shows the number of packets of chips sold from a vending machine over a 2 week period.

15, 17, 18, 18, 14, 16, 17, 6, 16, 18, 16, 16, 20, 18

Display this information in a dot plot.

c Explain why a dot plot is only suitable for discrete data.

12 Explain why it would not be suitable to represent each of the following data sets on a dot plot.

a A cricketers scores in a season were 4, 65, 82, 5, 19, 56, 23, 153.

b The heights of 8 bushes were 1.93 m, 1.76 m, 1.55 m, 1.86 m, 1.97 m, 1.13 m, 1.05 m, 2.06 m.

13 Tina wants to draw graphs to represent the following data:

- The average daily sales in her shop over a week.
- The most frequently sold items in her shop.
- The amounts of money spent by her customers.

Which of these would be most suitable to display on a radar chart? Explain your choice.

14 For each of the following data sets choose the most appropriate graph type.

a The number of car registrations each year in NSW.

b The average road wear of five different brands of tyre.

c A comparison of the different fuel types sold by a petrol station.

d The number of cars that enter a petrol station each hour over a 24 hour period.

15 Explain why categorical data is unsuitable to be represented by a line graph.

eBookplus
Digital doc
EXCEL Spreadsheet
doc-1563
Column graphs

INVESTIGATE: Choice of graph

For different statistical investigations, some graphs may be more suitable than others. Depending on the information that needs to be displayed, each graph has its strengths and weaknesses.

In the above exercise, you drew a dot plot, sector graph, line graph, column graph and radar chart. For each of these graph types describe:

a The advantages of this type of display.

b The disadvantages of this type of display.

c An example of when this type of display would be most suitable.

eBookplus
Digital doc
EXCEL Spreadsheet
doc-1564
Line graphs

eBookplus
Digital doc
EXCEL Spreadsheet doc-1565
Bar graphs

eBookplus
Digital doc
EXCEL Spreadsheet doc-1566
Sector graphs

INVESTIGATE: Producing graphs using technology

Each of the graph types described in this section can be drawn using a spreadsheet or a graphics calculator. From the *Maths Quest General Mathematics Preliminary Course* eBookPLUS, click on each of the spreadsheet icons in the margin to practise using a spreadsheet to produce a graph.

9C Statistical graphs

When displaying statistical, or quantitative data, the most useful graphs are a **frequency histogram** and frequency polygon.

A **histogram** is similar to a column graph with the following essential features.

- Gaps are never left between the columns, except for a half unit space before the first column.
- If the chart is coloured or shaded then it is done all in one colour. (The columns are essentially all representing different levels of the same thing.)
- Frequency is always plotted on the vertical axis.
- For ungrouped data, the horizontal scale is marked so that the data labels appear under the centre of each column. For grouped data, the horizontal scale is marked so that the class centre of each class appears under the centre of the column.

WORKED EXAMPLE 8

The table below shows the number of people living in each house in a street.

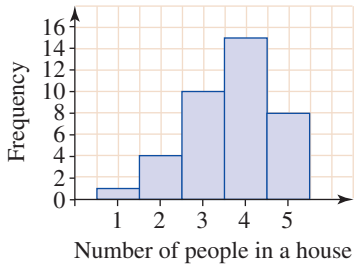
No. of people	Frequency
1	1
2	4
3	10
4	15
5	8

Show this information in a frequency histogram.

THINK

1. Draw a set of axes with the number of people living in a house on the horizontal axis and frequency on the vertical axis.
2. Draw the graph, leaving half a column width space before the first column.

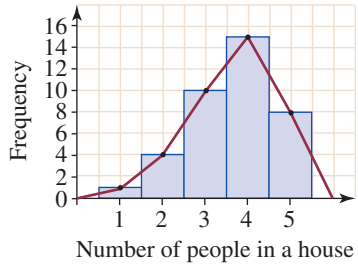
DRAW



A **frequency polygon** is a line graph that can be drawn by joining the centres of the tops of each column of the histogram. The **polygon** starts and finishes on the horizontal axis a half column width space from the group boundary of the first and last groups.

The figure at right shows the frequency polygon drawn on top of the histogram for the previous worked example.

It is common practice to draw the histogram and the polygon on the same set of axes.



WORKED EXAMPLE 9

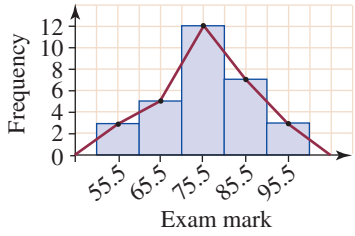
The frequency table below shows a class set of marks on an exam. Draw a frequency histogram and polygon on the same set of axes.

Mark	Class centre	Frequency
51–60	55.5	3
61–70	65.5	5
71–80	75.5	12
81–90	85.5	7
91–100	95.5	3

THINK

1. Draw a set of axes with the exam mark on the horizontal axis and frequency on the vertical axis. Show the class centres for the exam marks.
2. Draw the columns, leaving a half column width space before the first column.
3. Draw a line graph to the centre of each column.

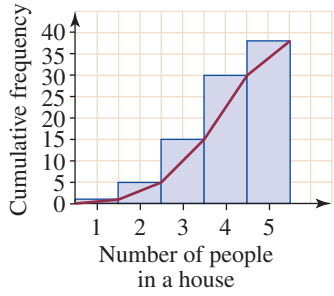
DRAW



For further analysis of statistical data it is useful to draw cumulative frequency graphs. The **cumulative frequency** is a progressive total of the frequency column. Consider the table from worked example 8.

No. of people	Frequency	Cumulative frequency
1	1	1
2	4	5
3	10	15
4	15	30
5	8	38

Once the cumulative frequency has been calculated, a **cumulative frequency histogram** and polygon (**ogive**) can be drawn. The cumulative frequency histogram is drawn without the half column space before the first column. The **cumulative frequency polygon** is drawn to the top right-hand corner of each column.



WORKED EXAMPLE 10

The frequency table below shows the heights of people in a basketball squad.

Height (cm)	Class centre	Frequency
170–174	172	3
175–179	177	6
180–184	182	12
185–189	187	10
190–194	192	8
195–199	197	1

a Add a cumulative frequency column to the table.
b Draw a frequency histogram and polygon.

THINK

a 1 Add a fourth column to the table.

2 Complete the column by keeping a running total of the frequencies.

b 1 Draw the axes with height on the horizontal axis and cumulative frequency on the vertical axis. Show the class centres for the height.

2 Draw the columns for the cumulative frequency histogram.

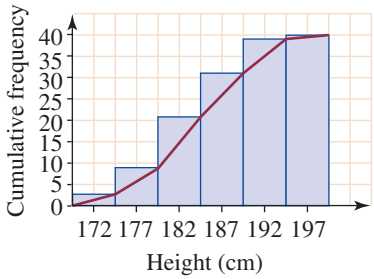
3 Draw the cumulative frequency polygon by joining the lines to the top right-hand corner of each column.

WRITE/DRAW

a

Height (cm)	Class centre	Frequency	Cumulative frequency
170–174	172	3	3
175–179	177	6	9
180–184	182	12	21
185–189	187	10	31
190–194	192	8	39
195–199	197	1	40

b



REMEMBER

1. A piece of numerical data is known as a score.
2. Quantitative data can be graphed using histograms and polygons.
3. When drawing histograms, always put frequency on the vertical axis and never leave gaps between columns.
4. If the histogram is illustrating ungrouped data, the data labels on the horizontal axis are placed under the centre of each column.
5. If the histogram is illustrating grouped data, the data labels on the horizontal axis (that is, the class centres) are placed under the centre of each column.
6. A polygon is a line graph that can be drawn by joining the centres of the tops of each column of the histogram.
7. The cumulative frequency is a running total of the frequency column. The cumulative frequency can also be graphed using a cumulative frequency histogram and polygon.

EXERCISE 9C

Statistical graphs

eBookplus
Digital doc
SkillSHEET 9.5
doc-1567
Producing a frequency table from a frequency histogram

1 WE8 A survey was done on young drivers taking the written test for their licence.
The number of mistakes each made was recorded and the results are shown in the frequency distribution table at right.
Show this information in a frequency histogram.

No. of mistakes (score)	No. of drivers (frequency)
0	5
1	8
2	13
3	4
4	3
5	1

eBookplus
Digital doc
EXCEL Spreadsheet
doc-1568
Histogram, frequency polygon

2 Each student in a class was asked the number of children in their family.
The results are shown in the frequency distribution table at right.
Show this information in a frequency histogram and polygon.

No. of children in a family	Frequency
1	3
2	5
3	8
4	4
5	2
6	1

eBookplus
Digital doc
EXCEL Spreadsheet
doc-1569
Histogram, frequency polygon (DIY)

3 The table below shows the age in years of the members of a surf club.

Age	No. of members
18	3
19	5
20	8
21	13
22	15
23	10
24	8
25	5

Show this information on a frequency polygon.

4 The label on a box of matches states that the average contents of a box is 50 matches. Quality control surveyed 50 boxes for the number of matches and the results are shown below.

48	50	50	51	50	49	53	52	48	51
50	50	51	49	48	53	52	50	49	49
49	50	50	51	53	52	54	47	50	49
48	49	47	53	49	52	50	51	50	50
50	48	47	50	51	49	50	49	52	51

a Put this information into a frequency table.
b Show the results on a frequency histogram and polygon.

5 WE9 The table below shows the length of 71 fish caught in a competition.

Length of fish (mm)	Class centre	Frequency
300–309	304.5	9
310–319	314.5	15
320–329	324.5	20
330–339	334.5	12
340–349	344.5	8
350–359	354.5	7

Show this information in a frequency histogram and polygon.

6 Sixty people were involved in a psychology experiment. The following frequency table shows the times taken for the 60 people to complete a puzzle for the experiment.

Time taken (seconds)	Class centre	Frequency
6–8		1
8–10		4
10–12		16
12–14		18
14–16		12
16–18		8
18–20		2

a Copy the frequency table and complete the class centre column.
b Show the information in a frequency histogram and polygon.

7 WE10 The following data show the number of registered cars normally kept at each of 30 households.

No. of cars	Frequency
0	2
1	8
2	11
3	6
4	2
5	1

a Copy the table and add a cumulative frequency column.
b Draw a cumulative frequency histogram and polygon.

8 The following table shows the number of jelly beans in each of 60 packets.

No. of jelly beans	Frequency
48	2
49	10
50	32
51	9
52	5
53	2

a Add a cumulative frequency column to the table.
b Draw a cumulative frequency histogram and polygon.

9 The following frequency table gives the number of oysters of different lengths from a tray in a marine farm.

Length (cm)	Class centre	Frequency	Cumulative frequency
4–5		6	
5–6		10	
6–7		60	
7–8		58	
8–9		8	
9–10		4	

a Copy and complete the table.
b Show the information in the form of a cumulative frequency histogram and polygon.

10 The following frequency table gives the results of testing the lives of 200 torch batteries.

Lifetime (hours)	Frequency
20–25	6
25–30	25
30–35	70
35–40	61
40–45	30
45–50	8

eBookplus
Digital doc
WorkSHEET 9.1
doc-1570

a Redraw the table, including a column for class centre and cumulative frequency.
b Draw a cumulative frequency histogram and polygon.

Further development

11 Explain why a width space needs to be left before the drawing of the first column of a frequency histogram.

12 A class of 30 students sat for a Mathematics test. Their results out of 100 are shown below.

68 72 58 45 69 92 38 51 70 65 69 73 52 76 48
69 73 41 42 73 80 50 60 49 65 94 88 85 53 60

a Use these results to copy and complete the frequency table below.

Class	Class centre	Tally	Frequency
30–39			
40–49			
50–59			
60–69			
70–79			
80–89			
90–99			

b Construct a histogram to display these data.

13 A farmer measures the heights of his tomato plants. The results, in metres, are shown below.

0.93 1.21 2.03 1.40 1.17 1.53 1.82 1.77 1.65 0.63 1.24 1.99
0.80 2.14 1.53 2.07 1.96 1.05 0.94 1.23 1.72 1.34 0.75 1.17
1.50 1.41 1.74 1.86 1.55 1.42 1.52 1.39 1.76 1.67 1.28 1.43 2.13

a Use the class groupings 0.6–<0.8, 0.8–<1.0, 1.0–<1.2, . . . etc. to complete a frequency distribution table for the data.
b Construct a histogram to display the data.

14 The following data give the times (in seconds) taken for athletes to complete a 100 m sprint.

12.2 12.0 11.9 12.0 12.6 11.7 11.4 11.0 10.9 11.7 11.2 11.8
12.2 12.0 12.7 12.9 11.3 11.2 12.8 12.4 11.7 10.8 13.3 11.7
11.6 11.7 12.2 12.7 13.0 12.2

a Construct a frequency table for the data using a class size of 5.
b Construct a histogram to display the data.

15 The frequency table below shows the lengths (in millimetres) of 77 flathead caught in a fishing competition.
a Copy the table and add a cumulative frequency column to it.
b Prepare an ogive of the data.

Length of fish (mm)	Frequency
300–310	9
310–320	15
320–330	20
330–340	12
340–350	8
350–360	7
360–370	4
370–380	2

16 The following frequency table shows the time taken (in seconds) for 60 people, involved in a psychology experiment, to complete a simple manipulative puzzle.

Time taken (s)	Frequency
6–8	1
8–10	4
10–12	15
12–14	18
14–16	12
16–18	8
18–20	2

a Copy the table and add a cumulative frequency column to it.
b Prepare an ogive of the data.

17 The table below shows the frequency of scores, arranged by class, from a basketball team's normal season.

Class	Class centre	Frequency	Cumulative frequency
1–10		5	
11–20		15	
21–30		29	
31–40		37	
41–50		11	

a Copy the frequency table above and complete the class centre column.
b Complete the cumulative frequency column.
c How many scores in the data set were above 30?
d How many scores in the data set were 40 or less?
e Is the data set an example of grouped or ungrouped data?
f Draw a frequency histogram for the data set.
g On your histogram, draw a polygon for this data set.
h Draw a cumulative frequency histogram for the data set.
i On your cumulative frequency histogram, draw a cumulative frequency polygon.
j Copy and complete: Another term for a cumulative frequency polygon is an _____.

9D Range and interquartile range

Once a set of scores has been collected, tabulated and graphed, we are ready to make some conclusions about the data.

The range and interquartile range are used to measure the spread of a set of scores.

The **range** is the difference between the highest and the lowest score.

Range = highest score – lowest score

WORKED EXAMPLE 11

There are 17 players in the squad for a State of Origin match. The number of State of Origin matches played by each member of the squad is shown below.

2 6 12 8 1 4 8 9 24 4 5 11 14 6 11 15 10

What is the range of this distribution?

THINK

1. The lowest number of matches played is 1.
2. The highest number of matches played is 24.
3. Calculate the range by subtracting the lowest score from the highest score.

WRITE

Lowest score = 1

Highest score = 24

Range = 24 – 1
= 23

A smaller range will usually represent a more consistent set of scores. Exceptions to this are when one or two scores are much higher or lower than most.

When we are calculating the range from a frequency distribution table, we find the highest and lowest score from the score column. We do not use any information from the frequency column in calculating the range. When the data are presented in grouped form, the range is found by taking the highest score from the highest class and the lowest score from the lowest class.

WORKED EXAMPLE 12

The frequency distribution table at right shows the heights of boys competing for a place on a basketball team. Find the range of these data.

Height	Frequency
170–175	3
175–180	6
180–185	12
185–190	10
190–195	8
195–200	1

THINK

1. The lowest score is at the bottom of the 170–175 class.
2. The highest score is at the top of the 195–200 class.
3. Range = highest score – lowest score.

WRITE

Lowest score = 170

Highest score = 200

Range = 200 – 170
= 30

In many cases, the range is not a good indicator of the overall spread of scores. Consider the two sets of scores below showing the wages of people in two small businesses.

A: \$240, \$240, \$240, \$245, \$250, \$250, \$260, \$800
B: \$180, \$200, \$240, \$290, \$350, \$400, \$500, \$600

The range for business A = \$800 – \$240 = \$560 and for business B = \$600 – \$180 = \$420

While the range for business A is greater, by looking at the wages in the two businesses, we can see that the wages in business B are generally more spread. The range uses only two scores in its calculation. The interquartile range is usually a better measure of dispersion (spread).

The **quartiles** are found by dividing the data into quarters. The **lower quartile** is the lowest 25% of scores, the **upper quartile** is the highest 25% of scores.

To calculate an interquartile range, we must first be able to calculate the **median**. To calculate the median, we must first arrange the scores in ascending order. The median is the middle score (if there is an odd number of scores) or the average of the two middle scores (if there is an even number of scores).

WORKED EXAMPLE 13

Calculate the median of:
a 2, 5, 8, 8, 8, 11, 12
b 45, 69, 69, 87, 88, 92, 99, 100.

THINK

a There are 7 scores so the median is the 4th score.

b There are 8 scores so the median is the average of the 4th score and the 5th score.

WRITE

a Median = 8

b $\text{Median} = \frac{87 + 88}{2}$
$= 87.5$

The **interquartile range** is the difference between the upper quartile and the lower quartile. To find the lower and upper quartiles we arrange the scores in ascending order. The lower quartile is $\frac{1}{4}$ of the way through the distribution and the upper quartile is $\frac{3}{4}$ of the way through the distribution.

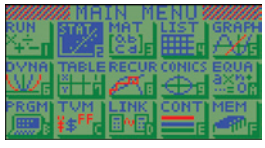
To find the interquartile range we follow the steps below.

1. Arrange the data in ascending order.
2. Divide the data into two halves by finding the median.
   (a) If there is an odd number of scores, the median score should not be included in either half of the scores.
   (b) If there is an even number of scores, the middle will be halfway between two scores and this will divide the data neatly into two sets.
3. The lower quartile will be the median of the lower half of the data.
4. The upper quartile will be the median of the upper half of the data.
5. The interquartile range will be the difference between the medians of the two halves of the data.

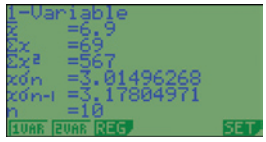
WORKED EXAMPLE 14

Find the interquartile range of the following data, which show the number of home runs scored in a series of baseball matches.

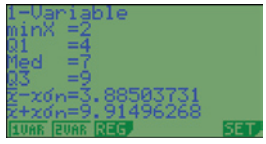
12, 9, 4, 6, 5, 8, 9, 4, 10, 2

THINK	WRITE
Method 1	
1 Write the data in ascending order.	2, 4, 4, 5, 6, 8, 9, 9, 10, 12
2 Divide the data into two equal halves.	2, 4, 4, 5, 6 8, 9, 9, 10, 12
3 The lower quartile is the median of the lower half.	Lower quartile = 4
4 The upper quartile is the median of the upper half.	Upper quartile = 9
5 The interquartile range is the upper quartile minus the lower quartile.	Interquartile range = 9 – 4 = 5
Method 2	
1 From the **MENU** select **STAT**.	
2 Delete any existing data from all lists and then enter the scores into **List 1**.	
3 Press F2 for **CALC** and then F6 for **SET**. Enter the settings as shown at right. **1Var Xlist:List 1** shows that the scores are stored in List 1. **1Var Freq:1** shows that each score in List 1 is an individual score with a frequency of 1.	

4. Press EXE to return to the previous screen and then press F1 for **1Var** and all the summary statistics will be displayed.



5. Scroll down using the arrow keys until you can see the median, which is equal to 7.



6. To find the range, we need to find the lowest and the highest score. On the previous screen you will see the lowest score denoted **MinX**. Scroll down further to find **MaxX**. The range is found by subtracting **MinX** from **MaxX**.

Range = 12 – 2
= 10

7. On this screen you will see the value of the upper quartile Q3 and the lower quartile Q1. To find the interquartile range, subtract Q1 from Q3.

IQR = 9 – 4
= 5

In most cases we are asked to find the interquartile range of a grouped distribution. This requires us to draw a cumulative frequency polygon and find the 25th and 75th percentile.

A percentile is a measure of where in a set of scores an individual score lies. For example, the 25th percentile has 25% of scores below it and 75% above it.

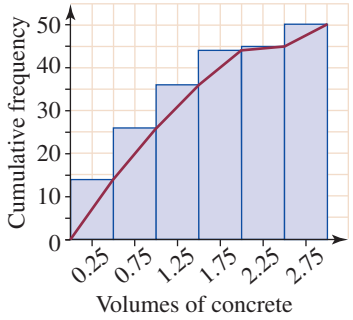
To find the interquartile range, draw a second vertical axis that shows the 25th, 50th and 75th percentile. A line is drawn from the 25th, 50th and 75th percentile to the ogive and then down to the horizontal axis. The value for the quartiles can then be calculated.

The median is the score that is found at the 50th percentile.

WORKED EXAMPLE 15

eBook*plus*
Tutorial
int-2325
Worked example 15

The cumulative frequency histogram and polygon at right shows the number of customers who order different volumes of concrete from a readymix concrete company during a day.
Find the:
a median
b interquartile range for this distribution.



THINK

a 1 Draw a vertical axis showing the percentiles.

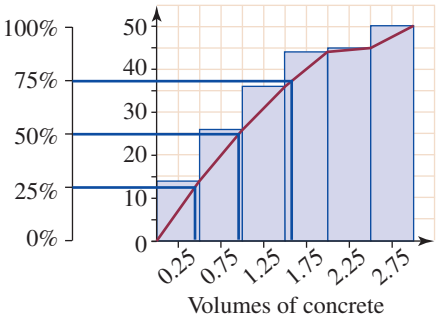
2 Draw a line for the 50th percentile to the ogive and estimate the median.

b 1 Draw a line for the 25th and 75th percentiles and estimate these values.

2 Calculate the interquartile range by subtracting the lower quartile from the upper quartile.

WRITE/DRAW

a



Median = 0.9

b Lower quartile = 0.4
Upper quartile = 1.6
Interquartile range = 1.6 – 0.4
= 1.2

A data set can also be divided into deciles. A **decile** is a band of 10% of all scores. The deciles can be calculated in the same way as the quartiles, using the appropriate percentage on the vertical scale. For example, the top decile would be found using the 90th percentile.

REMEMBER

1. The range and interquartile range, also called measures of dispersion, are used to measure the spread of a set of scores.
2. The range is calculated by subtracting the lowest score from the highest score.
3. A single outlying score can enlarge the range. The interquartile range is therefore a better measure of dispersion.
4. The interquartile range is found by subtracting the lower quartile from the upper quartile.
5. The lower and upper quartiles are found by dividing the scores into two equal halves. The median of the lower half is the lower quartile and the median of the upper half is the upper quartile.
6. The interquartile range is estimated from a grouped distribution by drawing a cumulative frequency histogram and polygon.

EXERCISE 9D

Range and interquartile range

eBookplus
Digital doc
SkillSHEET 9.6
doc-1571
Finding the median of a small data set

1 WE11 Find the range of each of the following sets of data.

a 2, 5, 4, 5, 7, 4, 3
b 103, 108, 111, 102, 111, 107, 110
c 2.5, 2.8, 3.4, 2.7, 2.6, 2.4, 2.9, 2.6, 2.5, 2.8
d 3.20, 3.90, 4.25, 7.29, 1.45, 2.77, 8.39
e 45, 23, 7, 47, 76, 89, 96, 48, 87, 76, 66

eBookplus
Digital doc
GC program — Casio
doc-1572
UV stats

eBookplus
Digital doc
GC program — TI
doc-1573
UV stats

2 WE12 Use the frequency distribution tables below to find the range for each of the following sets of scores.

a

Score	Frequency
1	2
2	6
3	12
4	10
5	7

b

Score	Frequency
38	23
39	46
40	52
41	62
42	42
43	45

c

Score	Frequency
89	12
90	25
91	36
92	34
93	11
94	9
95	4

3 For the grouped dispersions below, state the range.

a

Class	Frequency
51–60	2
61–70	8
71–80	15
81–90	7
91–100	1

b

Class	Frequency
150–155	12
155–160	25
160–165	38
165–170	47
170–175	39
175–180	20

c

Class	Frequency
40–43	48
44–47	112
48–51	254
52–55	297
56–59	199
60–63	84

4 The scores below show the number of points scored by two AFL teams over the first 10 games of the season.

Sydney:	110	95	74	136	48	168	120	85	99	65
Collingwood:	125	112	89	111	96	113	85	90	87	92

a Calculate the range of the scores for each team.
b Based on the results above, which team would you say is the more consistent?

5 Two machines are used to fill boxes with approximately 100 Smarties. A check is made on the operation of the two machines. Ten boxes filled by each machine have the number of Smarties in them counted. The results are shown below.

Machine A: 100, 99, 99, 101, 100, 101, 100, 100, 101, 108
Machine B: 98, 104, 96, 97, 103, 96, 102, 100, 97, 104

a What is the range in the number of Smarties from machine A?
b What is the range in the number of Smarties from machine B?
c Ralph is the quality control officer and he argues that machine A is more consistent in its distribution of Smarties. Explain why.

eBookplus
Digital doc
EXCEL Spreadsheet
doc-1574
Median

6 WE13 Find the median for each of the data sets below.

a 3, 4, 4, 5, 7, 9, 10
b 17, 20, 19, 25, 29, 27, 28, 25, 29
c 52, 55, 53, 53, 54, 55, 52, 53, 54, 52
d 12, 14, 15, 12, 14, 19, 17, 15, 18, 20
e 56, 75, 83, 47, 93, 35, 84, 83, 73, 20, 66, 90

eBookplus
Digital doc
EXCEL Spreadsheet
doc-1575
Median (DIY)

7 WE14 For each of the data sets in question **6**, calculate the interquartile range.

8 WE15 The frequency histogram and polygon at right displays the results of a survey of 50 drivers who were asked about the number of speeding fines they have received.

a Use the ogive to find the median of the distribution.
b Find the lower quartile.
c Find the upper quartile.
d Calculate the interquartile range.
e Calculate the top decile.

9 The frequency distribution table below shows the result of a survey of 90 households who were asked about the number of times they had been the victim of crime.

Score	Frequency
0	26
1	31
2	22
3	8
4	3

a Add a column for cumulative frequency to the table.
b Draw a cumulative frequency histogram and polygon.
c Use your graph to find the median of the distribution.
d Calculate the interquartile range.

10 MC For the frequency table below, what is the range?

Score	Frequency
25	14
26	12
27	19
28	25
29	19

A 4 **B** 5 **C** 6 **D** 17

11 MC Calculate the interquartile range of the following data.
17, 18, 18, 19, 20, 21, 21, 23, 25

A 3 **B** 4 **C** 5 **D** 8

12 MC The interquartile range is considered to be a better measure of the variability of a set of scores than the range because it:

- **A** takes into account more scores
- **B** is the difference between the upper and lower quartiles
- **C** is easier to calculate
- **D** is not affected by extreme values.

13 MC The distribution below shows the ranges in the heights of 25 members of a football squad.

Height (cm)	Class centre	Frequency	Cumulative frequency
140–149	144.5	2	2
150–159	154.5	5	7
160–169	164.5	10	17
170–179	174.5	7	24
180–189	184.5	1	25

Which of the statements below is correct?

- **A** The range of the distribution is 40.
- **B** The range of the distribution is 49.
- **C** The range of the distribution is 9.
- **D** The interquartile range can only be estimated using a cumulative frequency polygon.

14 The frequency distribution table below shows the marks obtained by a group of people on an IQ test.

IQ score	Frequency
75–85	12
85–95	25
95–105	50
105–115	24
115–125	13

- **a** Redraw the frequency distribution table to include columns for class centre and cumulative frequency.
- **b** Draw a cumulative frequency histogram and polygon.
- **c** Find the range.
- **d** Use the graph to estimate the interquartile range.

15 The following frequency distribution table shows the distribution of daily maximum temperatures during the course of a full year.

Maximum temperature (°C)	Number of days
0–5	4
5–10	22
10–15	95
15–20	124
20–25	94
25–30	19
30–35	5
35–40	2

a Add a cumulative frequency column to the table.
b Draw an ogive of the data.
c Find the upper and lower quartiles of the data and calculate the interquartile range.
d Use the ogive to find the median (50th percentile of the data).
e Find the bottom decile of the data (the scores between which the lowest 10% of scores lie).

16 The following data give the number of fruit that have formed on each of 30 trees in an orchard.

45	48	52	36	38	72	36	74	56	46
81	73	46	48	44	39	52	58	57	65
60	53	54	58	41	44	47	76	68	55

a Complete a frequency distribution table for the data.
b Draw an ogive of the data.
c Use the ogive to find the median, lower quartile and upper quartile of the data.
d Find the interquartile range of the data.

Further development

17 The salaries of the 40 employees of a small manufacturing company are represented by the accompanying frequency table.

Salary (× $1000)	Frequency
15–20	6
20–25	12
25–30	8
30–35	7
35–40	5
40–45	1
45–50	1

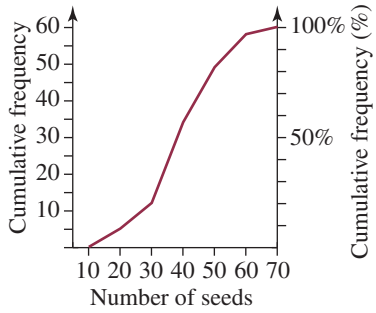
a Copy the table and add a cumulative frequency column to it.
b Prepare an ogive of the data.
c How many employees are earning less than $22 000?
d How many employees are earning less than $31 000?
e Find the 75th percentile of the data and write a sentence explaining what it means.
f Find the 50th percentile of the data and write a sentence explaining what it means.
g Find the 25th percentile of the data and write a sentence explaining what it means.
h The management decides to award pay rises to its highest earning employees. The top 10% of employees will all get a pay rise. How much salary would an employee need before qualifying for a pay rise?

18 A manufacturer of surf clothing needs to know how many clothes of different sizes to produce. The manager organises a survey of young people which provides the following data:

Waist size (cm)	Frequency
70–75	13
75–80	28
80–85	46
85–90	30
90–95	17
95–100	8
100–105	7
105–110	1

a Copy the table and add a cumulative frequency column to it.
b Prepare an ogive of the data.
c How many young people had a waist size of less than 82 cm?
d How many young people had a waist size of greater than 94 cm?
e Find the 90th percentile of the data and write a sentence explaining what it means.
f Find the 50th percentile of the data and write a sentence explaining what it means.
g The manager decides that production costs can be minimised by only making garments fitting sizes between 78 cm and 100 cm. What percentage of the population will not be catered for by this manufacturer?

19 A biologist who counts the number of seeds produced in each of 60 pumpkins presents his findings on the ogive below.



a How many pumpkins contained 30 or fewer seeds?
b How many pumpkins contained more than 50 seeds?
c What percentage of pumpkins had fewer than 45 seeds?
d What percentage of pumpkins had fewer than 20 seeds?
e Find the 90th percentile of the data and write a sentence explaining what it means.
f Find the 75th percentile of the data and write a sentence explaining what it means.
g Find the 50th percentile of the data and write a sentence explaining what it means.
h The worst 20% of pumpkins (those with the fewest seeds) are to be kept aside for further investigation. Find the maximum number of seeds for any pumpkin in this group.

20 A time trial is a race in which each competitor rides separately, racing 'against the clock'. The following are the times (in seconds) of 20 competitors in a 1 km cycling time trial.

75 72 68 78 75 68 77 80 85 82
73 75 82 90 92 75 73 72 70 83

a Copy and complete the frequency table below.

Time (s)	Frequency	Cumulative frequency
65–70		
70–75		
75–80		
80–85		
85–90		
90–95		

b Prepare an ogive of the data.
c How many riders finished with a time of 82 seconds or better?
d Find the 90th percentile and write a sentence which explains what it means.
e The top 20% of riders are to be selected for a special training squad. What time would be needed to qualify for the squad?

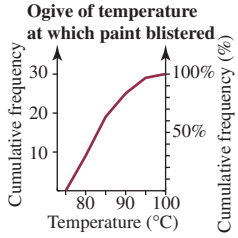
21 The following data, collected from a maternity hospital, gives the birth weights (in kg) of 30 babies.

3.7	3.2	3.8	4.1	2.9	3.3	3.6	3.1	3.6	3.9
4.2	2.5	2.7	3.9	3.6	3.2	3.0	2.9	3.4	3.0
3.1	2.8	2.9	3.2	3.1	3.8	3.9	3.3	4.4	3.4

a Using classes 2.4–2.8, 2.8–3.2, . . . etc. complete a frequency distribution table.
b Prepare an ogive of the data.
c What percentage of babies had a birth weight of 3.8 kg or less?
d Find the 50th percentile and explain what it means.
e Babies that weigh 2.6 kg or less are given special attention by medical staff. What percentage of babies are given special attention?

22 The ogive below shows the temperatures (in degrees Celsius) at which paint starts to blister.

a How many paint samples blistered at a temperature of 85 °C or less?
b What percentage of the samples blistered at 85 °C or less?
c What percentage of the samples could withstand a temperature of 88 °C?
d Find the 90th percentile and explain what it means.
e The manufacturer wishes to guarantee that the paint will not blister at high temperatures. What is the highest temperature at which the paint could be rated, if management wants no more than 20% of the paint returned because of blistering?



9E Stem-and-leaf plots

As an alternative to a frequency table, a **stem-and-leaf plot** may be used to group and summarise data.

A stem is made using the first part of each piece of data. The second part of each piece of data forms the leaves. Consider the case below.

The following data show the mass (in kg) of 20 possums trapped, weighed then released by a wildlife researcher.

1.8	0.9	0.7	1.4	1.6	2.1	2.7	2.2	1.8	2.3
2.3	1.5	1.1	2.2	3.0	2.5	2.7	3.2	1.9	1.7

The stem is made from the whole number part of the mass and the leaves are the decimal part. The first piece of data was 1.8 kg. The stem of this number could be considered to be 1 and the leaf 0.8. The second piece of data was 0.9. It has a stem of 0 and a leaf of 0.9.

To compose a stem-and-leaf plot for these data, rule a vertical column of stems then enter the leaf of each piece of data in a neat row beside the appropriate stem. The first row of the stem-and-leaf plot records all data from 0.0 to 0.9. The second row records data from 1.0 to 1.9 etc. Attach a key to the plot to show the reader the meaning of each entry.

It is convention to assemble the data in order of size, so this stem-and-leaf plot should be written in such a way that the numbers in each row of 'leafs' are in ascending order.

Key: 0 | 7 = 0.7 kg

Stem	*Leaf*
0	7 9
1	1 4 5 6 7 8 8 9
2	1 2 2 3 3 5 7 7
3	0 2

When preparing a stem-and-leaf plot, it is important to try to keep the numbers in neat vertical columns because a neat plot gives the reader an idea of the distribution of scores. The plot itself looks a bit like a histogram turned on its side.

WORKED EXAMPLE 16

The information below shows the mass, in kilograms, of twenty 16-year-old boys.

65 45 56 57 58 54 61 72 70 69
61 58 49 52 64 71 66 65 66 60

Show this information in a stem-and-leaf plot.

THINK

1. Make the 'tens' the stem and the 'units' the leaves.
2. Write a key.
3. Complete the plot.

WRITE

Key: 5 | 6 = 56 kg

Stem	*Leaf*
4	5 9
5	2 4 6 7 8 8
6	0 1 1 4 5 5 6 6 9
7	0 1 2

It is also useful to be able to represent data with a class size of 5. This could be done for the stem-and-leaf plot on the opposite page by choosing stems 0*, 1, 1*, 2, 2*, 3, where the class with stem 1 contains all the data from 1.0 to 1.4 and stem 1* contains the data from 1.5 to 1.9 etc. If stems are split in this way, it is a good idea to include two entries in the key. The stem-and-leaf plot for the 'possum' data would appear as follows.

Key: 1 | 1 = 1.1 kg 1* | 5 = 1.5

Stem	*Leaf*
0*	7 9
1	1 4
1*	5 6 7 8 8 9
2	1 2 2 3 3
2*	5 7 7
3	0 2

A stem-and-leaf plot has the following advantages over a frequency distribution table.

- **The plot itself gives a graphical representation of the spread of data. (It is rather like a histogram turned on its side.)**
- **All the original data are retained, so there is no loss of accuracy when calculating statistics such as the mean and standard deviation. In a grouped frequency distribution table some generalisations are made when these values are calculated.**

WORKED EXAMPLE 17

eBookplus
Tutorial
int-2326
Worked example 17

The following data give the length of gestation in days for 24 mothers. Prepare a stem-and-leaf plot of the data using a class size of 5.

280	287	285	276	266	292
288	273	295	279	284	271
292	288	279	281	270	278
281	292	268	282	275	281

THINK

1. A class size of 5 is required. The smallest piece of data is 266 and the largest is 295 so make the stems: 26*, 27, 27*, 28, 28*, 29, 29*. The key should give a clear indication of the meaning of each entry.
2. Enter the data piece by piece. Enter the leaves in pencil at first so that they can be rearranged into order of size. Check that 24 pieces of data have been entered.
3. Now arrange the leaves in order of size.

WRITE

Key: 26* | 6 = 266
27 | 0 = 270

Stem	*Leaf*
26*	6 8
27*	0 1 3
27*	5 6 8 9 9
28*	0 1 1 1 2 4
28*	5 7 8 8
29*	2 2 2
29*	5

Since all the original data are recorded on the stem-and-leaf plot and are conveniently arranged in order of size, the plot can be used to locate the upper and lower quartiles and the median.

As was discussed earlier in the chapter:

- the median is the middle score or the average of the two middle scores
- the lower quartile is the median of the lower half of the data
- the upper quartile is the median of the upper half of the data.

Using the 'possum' weight data as an example:
Key: 0 | 7 = 0.7 kg

Stem	*Leaf*
0	7 9
1	1 4 5 6 7 8 8 9
2	1 2 2 3 3 5 7 7
3	0 2

There were 20 pieces of data so the median is the average of the 10th and 11th scores. Counting each score as it appeared in the stem-and-leaf plot, we can see that the 10th score is the number 1.9 and the 11th score is the number 2.1.

$$\text{Median} = \frac{1.9 + 2.1}{2}$$
$$= 2.0$$

The median divides the data into two halves.

The lower quartile is the median of the lower half, which has ten scores in it. So the position of the lower quartile is given by the average of the 5th and 6th scores. The 5th score is the number 1.5. The 6th score is the number 1.6.

$$\text{The lower quartile} = \frac{1.5 + 1.6}{2}$$
$$= 1.55$$

The upper quartile is the median of the upper half, which also has ten scores in it. The 5th score in this half is the number 2.3. The 6th score is the number 2.5.

$$\text{The upper quartile} = \frac{2.3 + 2.5}{2}$$
$$= 2.4$$

WORKED EXAMPLE 18

Find the interquartile range of the data presented in the following stem-and-leaf plot.
Key: 15 | 7 = 157 kg

Stem	*Leaf*
15	**4 8 8**
16	**1 3 3 6 8**
17	**0 0 1 4 7 9 9 9**
18	**1 2 3 3 5 7 8 8 9**
19	**2 7 8**
20	**0 2**

THINK

1. There are 30 scores and so the median will be the average of the 15th and 16th scores.
2. There are 15 scores in each half and so the lower and upper quartiles will be the 8th score in each half.
3. The interquartile range is the difference between the upper and lower quartiles.

WRITE

$$\text{Median} = \frac{179 + 179}{2}$$
$$= 179$$

The lower quartile = 168
The upper quartile = 188

Interquartile range = upper quartile – lower quartile
= 188 – 168
= 20

REMEMBER

When presenting the stem-and-leaf plots, observe the following points.

1. Always include a key to assist in the interpretation of the plot.
2. Choose a suitable class size. A class size of 5 is possible by using * notation on class stems.
3. After initially recording each score, rearrange the leaves so that they appear in ascending order.

EXERCISE 9E

Stem-and-leaf plots

eBookplus

Digital doc
SkillSHEET 9.7
doc-1576
Presenting data as a stem-and-leaf plot

1 WE16 The data below give the number of errors made each week by 20 machine operators. Prepare a stem-and-leaf diagram of the data using stems of 0, 10, 20, etc.

6 15 20 25 28 18 32 43 52 27
17 26 38 31 26 29 32 46 13 20

2 The data below give the time taken for each of 40 runners on a 10 km fun run. Prepare a stem-and-leaf diagram for the data using a class size of 10 minutes.

36 42 52 38 47 59 72 68 57 82
66 75 45 42 55 38 42 46 48 39
42 58 40 41 47 53 68 43 39 48
71 42 50 46 40 52 37 54 48 52

3 WE17 The typing speed of 30 word processors is recorded below. Prepare a stem-and-leaf diagram of the data using a class size of 5.

96 102 92 96 95 102 95 115 110 108
88 86 107 111 107 108 103 121 107 96
124 95 98 102 108 112 120 99 121 130

4 Twenty transistors are tested by applying increasing voltage until they are destroyed. The maximum voltage that each could withstand is recorded below. Prepare a stem-and-leaf plot of the data using a class size of 0.5.

14.8 15.2 13.8 14.0 14.8 15.7 15.5 15.6 14.7 14.3
14.6 15.2 15.9 15.1 14.3 14.6 13.9 14.7 14.5 14.2

5 WE18 The stem-and-leaf plot at right gives the exact mass of 24 packets of biscuits. Find the interquartile range of the data.

Key: 248 | 4 = 248.4 g

Stem	*Leaf*
248	4 7 8
249	2 3 6 6
250	0 0 1 1 6 9 9
251	1 5 5 5 6 7
252	1 5 8
253	0

6 The time taken for a test vehicle to accelerate from 0 to 100 km/h is recorded during 24 trials. The results are represented by the stem-and-leaf plot at right.

a Find the median of the data.
b Find the upper and lower quartiles of the data.
c Find the interquartile range of the data.

Key: 7 | 2 = 7.2 s
7* | 6 = 7.6 s

Stem	*Leaf*
7	2 4 4
7*	5 5 7 9
8	0 0 1 2 4 4 4
8*	5 5 6 8 9
9	2 2 3
9*	0 2

Questions **7** to **10** refer to the stem-and-leaf plot at right.

Key: 12 | 1 = 1210
12* | 5 = 1250

Stem	*Leaf*
12	1 2 4
12*	5 7 7 9 9
13	0 1 1 2 3 4 4
13*	5 6 6 7 9 9
14	0 2 3 4
14*	0 1

7 MC The class size used in the stem-and-leaf plot is:
A 1 **B** 10 **C** 33 **D** 50

8 MC The number of scores that have been recorded is:
A 27 **B** 33 **C** 1210 **D** 1410

9 MC The median of the data is:
A 13.4 **B** 14 **C** 1335 **D** 1340

10 MC The interquartile range of the data is:
A 14 **B** 100 **C** 1290 **D** 1390

11 The maximum hand spans (in cm) of 20 male concert pianists is recorded as follows:

23.6 20.2 22.8 21.4 25.1 24.8 23.2 21.6 20.7 23.6
22.8 24.6 21.8 22.8 23.1 24.6 21.7 24.7 22.2 23.0

a Complete a stem-and-leaf plot to represent the data.
b Find the median of the data.
c Find the upper and lower quartile of the data.
d Find the interquartile range of the data.

12 The heights (in cm) of a sample of 30 plants are recorded as follows.

93 88 94 99 91 85 126 107 110 111
98 96 117 101 97 92 101 132 103 82
114 84 96 103 108 115 90 110 126 85

a Complete a stem-and-leaf plot to represent the data.
b Find the median of the data.
c Find the upper and lower quartile of the data.
d Find the interquartile range of the data.

Further development

13 The data below gives the head circumference (to the nearest centimetre) of 16 four-year-old girls.

48 49 47 52 51 50 49 48 50 50 53 52 43 47 49 50

a Draw a stem-and-leaf plot of the data using stems 4 and 5.

b Draw the stem-and-leaf plot with the stems 4 and 5 split into halve.

c Jessica decides to draw the stem-and-leaf plot in fifths, which is five rows for each stem, with 0's and 1's on the first row, 2's and 3's on the second row and so on. Complete the stem-and-leaf plot using this method.

14 A random sample of 20 screws is taken and the length of each is recorded to the nearest millimetre (below).

23 15 18 17 17 19 22 19 20 16
20 21 19 23 17 19 21 23 20 21

Construct a stem-and-leaf plot for screw length using:

a the stems 1 and 2

b the stems 1 and 2 split into halves

c the stems 1 and 2 split into fifths.

15 The number of seconds for which 12 Grade 2 children can hold their breath under water is given below.

8.2 9.2 8.1 8.5 9.3 8.9 8.9 9.5 8.9 9.0 9.1 9.7

Construct a stem-and-leaf plot for holding breath using:

a the stems 8 and 9

b the stems 8 and 9 split into halves

c the stems 8 and 9 split into fifths.

16 Explain the advantages of constructing a stem-and-leaf plot in either halves or fifths.

17 Explain why a stem-and-leaf plot can only be split into halves and fifths; that is, it cannot be split into quarters.

9F Five-number summaries

Once the median and quartiles have been calculated, we are able to summarise a data set using five numbers. This **five-number summary** consists of:

- **lower extreme** — the lowest score in the data set
- lower quartile — the score at the 25th percentile
- median — the middle score
- upper quartile — the score at the 75th percentile
- **upper extreme** — the highest score in the data set.

WORKED EXAMPLE 19

For the set of scores below, develop a five-number summary.
12 15 46 9 36 85 73 29 64 50

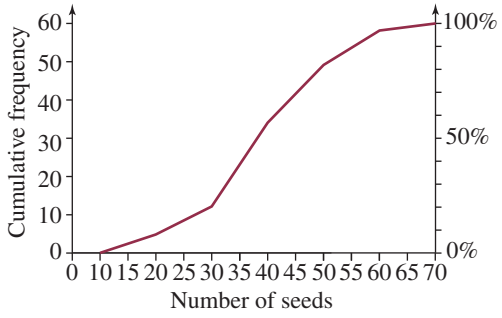
THINK	WRITE
1 Re-write the list in ascending order.	9 12 15 29 36 46 50 64 73 85
2 Write the lowest score.	Lower extreme = 9
3 Calculate the lower quartile.	Lower quartile = 15

4	Calculate the median.	$\text{Median} = \frac{36 + 46}{2}$ $= 41$
5	Calculate the upper quartile.	Upper quartile = 64
6	Write the upper extreme.	Upper extreme = 85 Five-number summary = 9, 15, 41, 64, 85

In most cases you will need to calculate the five-number summary from an ogive.

WORKED EXAMPLE 20

The ogive below shows the number of seeds found in each of 60 pumpkins.

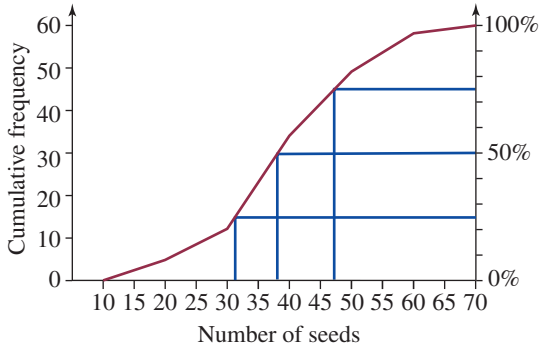


Use the ogive to develop a five-number summary.

THINK

1 Draw the 25th, 50th and 75th percentiles on the ogive.

WRITE/DRAW

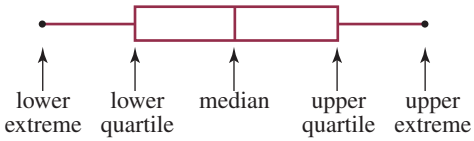


2	Write the lower extreme.	Lower extreme = 10
3	Use the ogive to estimate the lower quartile.	Lower quartile = 31
4	Use the ogive to estimate the median.	Median = 38
5	Use the ogive to estimate the upper quartile.	Upper quartile = 47
6	State the upper extreme.	Upper extreme = 70 Five-number summary = 10, 31, 38, 47, 70

eBookplus
Interactivity
int-0802
Boxplot and five-number summary

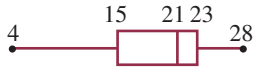
Once a five-number summary has been developed, it can be graphed using a **box-and-whisker plot**, a powerful way to display the spread of the data.

The box-and-whisker plot consists of a central divided box with attached whiskers. The box spans the interquartile range, the vertical line inside the box marks the median and the whiskers indicate the range.

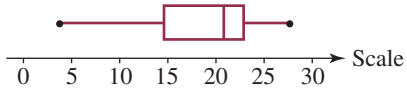


eBookplus
Interactivity
int-2401
Box-and-whisker plot

Box-and-whisker plots are always drawn to scale. This can be drawn with the five-number summary attached as labels:

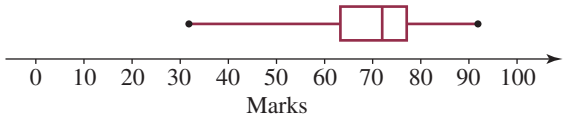


or with a scale presented alongside the box-and-whisker plot.



WORKED EXAMPLE 21

The box-and-whisker plot drawn below shows the marks achieved by a class on their end of year exam.



a State the median.
b Find the interquartile range.
c What was the highest mark in the class?

THINK	WRITE
a The mark in the box shows the median (72).	a Median = 72
b 1 The lower end of the box shows the lower quartile (63).	b Lower quartile = 63
2 The upper end of the box shows the upper quartile (77).	Upper quartile = 77
3 Subtract the lower quartile from the upper quartile.	Interquartile range = 77 − 63 = 14
c The top end of the whisker gives the top mark (92).	c Top mark = 92

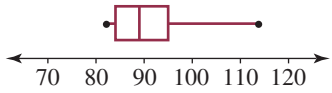
WORKED EXAMPLE 22

After analysing the speed of motorists through a particular intersection, the following five-number summary was developed.
The lowest score is 82.
The lower quartile is 84.
The median is 89.
The upper quartile is 95.
The highest score is 114.
Show this information in a box-and-whisker plot.

THINK

1. Draw a scale from 70 to 120 using 1 cm = 10 km/h.
2. Draw the box from 84 to 95.
3. Mark the median at 89.
4. Draw the whiskers to 82 and 114.

DRAW



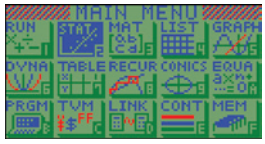
WORKED EXAMPLE 23

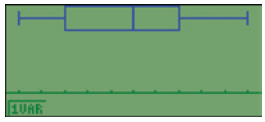
The number of home runs scored in a series of baseball games are: 12, 9, 4, 6, 5, 8, 9, 4, 10 and 2. use a graphics calculator to draw the box plot representing this data.

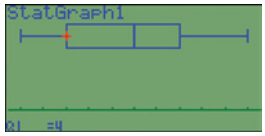
THINK

1. From the **MENU** select **STAT**.
2. Delete any existing data from all lists and then enter the scores into **List 1**.
3. Press F1 for **GRPH**, then F6 for **SET**. Use the arrow keys to highlight **Graph Type** and press F2 for Box, which will make **MedBox** appear in the display. Make sure that **Xlist** is set to **List 1** and **Frequency** is set to **1** as shown at right.
4. Press EXIT to return to the previous screen, then F1 for **GPH1**. The box-and-whisker plot then appears on the screen.
5. Press SHIFT F1 for **Trace**. You can then use the left and right arrow keys to display the values of the five-number summary used to draw the box-and-whisker plot.

DISPLAY







REMEMBER

1. A five-number summary is a summary set for a distribution.
2. The five numbers used in a five-number summary are the lower extreme, lower quartile, median, upper quartile and upper extreme.
3. A box-and-whisker plot can be used to graph a five-number summary.
4. The box is used to show the interquartile range and the median is marked with a line in the box. The whiskers then extend to show the range of the data set.

EXERCISE

9F Five-number summaries

eBookplus

Digital doc
GC program — Casio
doc-1572
UV stats

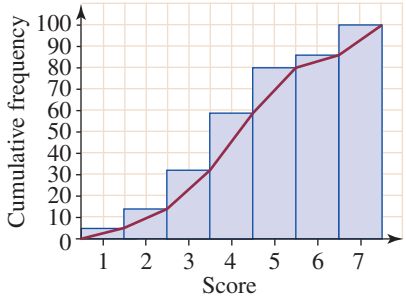
eBookplus

Digital doc
GC program — TI
doc-1573
UV stats

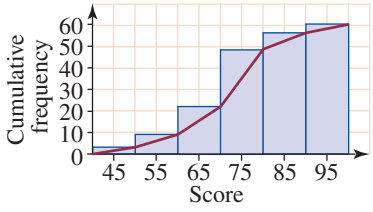
1 WE19 Write a five-number summary for the data set below.
15 17 16 8 25 18 20 15 17 14

2 For each of the data sets below, write a five-number summary.
a 23 45 92 80 84 83 43 83
b 2 6 4 2 5 7 1
c 60 75 29 38 69 63 45 20 29 93 8 29 93

3 WE20 A cumulative frequency histogram and polygon is shown at right. Write a five-number summary of the data set.



4 A cumulative frequency histogram and polygon is shown at right. Write a five-number summary of the data set.

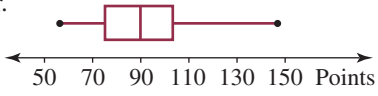


5 WE21 From the five-number summary 6, 11, 13, 16, 32 find:
a the median
b the interquartile range
c the range.

6 From the five-number summary 101, 119, 122, 125, 128 find:
a the median
b the interquartile range
c the range.

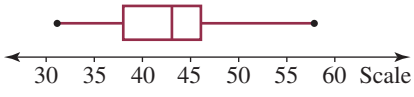
7 WE22 A five-number summary is given below.
Lower extreme = 39.2
Lower quartile = 46.5
Median = 49.0
Upper quartile = 52.3
Upper extreme = 57.8
Draw a box-and-whisker plot of the data.

8 The box-and-whisker plot below shows the distribution of final points scored by a football team over a season's roster.



a What was the team's greatest points score?
b What was the team's smallest points score?
c What was the team's median points score?
d What was the range of points scored?
e What was the interquartile range of points scored?

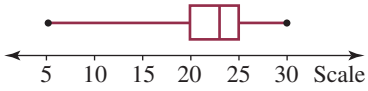
9 The box-and-whisker plot below shows the distribution of data formed by counting the number of honey bears in each of a large sample of packs.



In any pack, what was:

a the largest number of honey bears?
b the smallest number of honey bears?
c the median number of honey bears?
d the range of numbers of honey bears?
e the interquartile range of honey bears?

Questions **10**, **11** and **12** refer to the box-and-whisker plot drawn below.



10 MC The median of the data is:

A 20 **B** 23 **C** 35 **D** 31

11 MC The interquartile range of the data is:

A 23 **B** 26 **C** 5 **D** 20 to 25

12 MC Which of the following is *not* true of the data represented by the box-and-whisker plot?

A One-quarter of the scores is between 5 and 20.
B One-half of the scores is between 20 and 25.
C The lowest quarter of the data is spread over a wide range.
D Most of the data are contained between the scores of 5 and 20.

13 WE23 The data below show the number of sales made each day by a salesperson over a fortnight.

25, 31, 28, 43, 37, 43, 22, 45, 48, 33

a Write a five-number summary of the data.
b Draw a box-and-whisker plot of the data.

14 The data below show monthly rainfall in millimetres.

Jan.	Feb.	Mar.	Apr.	May	June	July	Aug.	Sept.	Oct.	Nov.	Dec.
10	12	21	23	39	22	15	11	22	37	45	30

a Provide a five-number summary of the data.
b Draw a box-and-whisker plot of the data.

15 The following data detail the number of hamburgers sold by a fast food outlet every day over a 4-week period.

Mon.	Tues.	Wed.	Thur.	Fri.	Sat.	Sun.
125	144	132	148	187	172	181
134	157	152	126	155	183	188
131	121	165	129	143	182	181
152	163	150	148	152	179	181

a Draw a frequency table of the data. (Use a class size of 10.)
b Draw a cumulative frequency histogram and polygon for the data.
c Use the ogive to find approximations for the:
  i median
  ii lower quartile
  iii upper quartile.
d Draw a box-and-whisker plot of the data.

16 The following data show the ages of 30 mothers upon the birth of their first baby.

22	21	18	33	17	23	22	24
24	20	25	29	32	18	19	22
23	24	28	20	31	22	19	17
23	48	25	18	23	20		

a Prepare a frequency table for the data. (Use a class size of 5.)
b Draw an ogive for the data.
c Draw a box-and-whisker plot of the data.
d Describe the distribution in words. What does the distribution say about the age that mothers have their first baby?

eBookplus
Digital doc
WorkSHEET 9.2
doc-1579

Further development

17 Employees on a factory assembly line are timed as they assemble a particular product. The results are below (in minutes).

18	40	31	37	18	46	27	20	54	35	14	48	23	38
19	39	18	28	37	24	32	16	43	39	37	46	24	35

a Prepare a five number summary of the data by first putting the data in a stem-and-leaf plot.
b Find the range.
c Find the interquartile range.
d If the slowest 25% of workers were to be dismissed, what would be the cut off time used?
e If you took 21 minutes to assemble an item, write a sentence to convince your employer that you are a valuable employee.

18 Two classes sat for a Mathematics test. Their results have been summarised by the five number summary below.

Class A: 25 40 52 75 95
Class B: 20 35 56 75 85

a Find the range of marks for each class.
b What is the median for each class?
c Calculate the interquartile range for each class.
d Which class has been the most consistent? Explain your answer.

19 The stem plot below details the age of 25 offenders who were caught during random breath testing.

Stem	*Leaf*
1	8 8 9 9 9
2	0 0 0 1 1 3 4 6 9
3	0 1 2 7
4	2 5
5	3 6 8
6	6
7	4

a Prepare a five number summary of the data.
b Draw a boxplot of the data.

20 Explain why a five number summary and boxplot is a useful way of displaying data.

21 Explain what the boxplot below tells you about the distribution of the data it represents.

SUMMARY

Frequency tables

- A frequency table is used to display a set of data in table form.
- Ungrouped data are placed in a table and every score is displayed in the table.
- Grouped data are used for continuous data or when the scores are spread over a large range. It is best to group data to create five to six classes. If the data are grouped, the table should display a column for class centre.

Types of graphs

- A dot plot is used to display a set of scores on a scale.
- A sector graph (pie chart) is used for a display that allows comparison of categorical data.
- A line graph is used to show the way a quantity changes over time.
- A bar or column graph is used to show quantities associated with categorical data.
- A radar chart is a type of line graph that shows the way in which a quantity changes over time. It is most appropriate for a period of time that repeats.

Statistical graphs

- Quantitative data are best displayed by a frequency histogram and polygon.
- A frequency histogram is a column graph that is drawn with a 0.5 unit (half column) space before the first column and no other spaces between the columns.
- A frequency polygon is drawn as a line graph from the corner of the axes to the centre of each column.
- A cumulative frequency histogram and polygon graphs the cumulative frequency. There is no space before the first column and the polygon is drawn from the corner of the axes to the top right-hand corner of each column.

Range and interquartile range

- The range is the difference between the highest score and the lowest score.
- The interquartile range is the difference between the score at the 25th percentile and the 75th percentile.
- The median is the score in the middle of the distribution (50th percentile).
- The median, lower quartile and upper quartile can be calculated by using an ogive (cumulative frequency polygon).

Five-number summaries

- A five-number summary of a data set is the lower extreme, lower quartile, median, upper quartile and upper extreme.
- A five-number summary can be graphed using a box-and-whisker plot.
- A box-and-whisker plot shows the spread of a data set on a scale.

CHAPTER REVIEW

MULTIPLE CHOICE

1 MC The frequency table below shows the marks achieved on a test by a group of students.

Score	Frequency
15	12
16	15
17	10
18	6
19	3
20	2

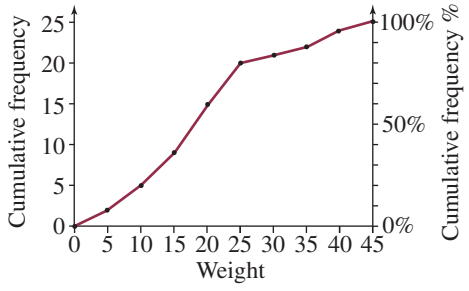
How many students received a mark higher than 17?

A 3 B 10 C 11 D 18

2 MC Which of the following would be greatly affected by the addition of an extreme score to the data set?

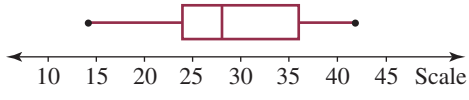
A the median
B the range
C the interquartile range
D all would be greatly affected

3 MC For the cumulative frequency polygon below, the interquartile range of the data is:



A 12 B 17
C 24 D 12 to 24

Questions **4** and **5** refer to the box-and-whisker plot shown below.



4 MC The upper quartile of the data is:

A 24 B 28 C 36 D 42

5 MC The interquartile range is:

A 12 B 28
C 14 to 42 D 24 to 36

SHORT ANSWER

1 A cricketer keeps track of the number of wickets he took in each innings in which he bowled during a season. The results are shown below.

0 2 3 1 0 6 4 1 1 2 0 0 4
3 2 2 3 1 0 1 1 1 2 0 1

Show this information in a frequency table.

2 The following data give the amount of cut meat (in kg) obtained from 20 lambs.

4.5 6.2 5.8 4.7 4.0 3.9 6.2 6.8 5.5 6.1
5.9 5.8 5.0 4.3 4.0 4.6 4.8 5.3 4.2 4.8

Show these data in a frequency table using a class size of 0.5 kg.

3 A survey is taken about the television stations being watched at 7:30 pm on a Monday night.

ABC – 27 SBS – 5 Channel 7 – 48
Channel 9 – 72 Channel 10 – 28

Show this information using a sector graph.

4 The number of absences from a school is taken over a week.

Monday – 43 Tuesday – 55 Wednesday – 34
Thursday – 45 Friday – 63

Show this information using a radar chart.

5 The table below shows the number of sales made each day over a month in a car yard.

Number of sales	Frequency
0	2
1	7
2	12
3	6
4	2
5	0
6	1

Show this information in a frequency histogram and polygon.

6 The frequency table below shows the crowds at football matches for a team over a season.

Class	Class centre	Frequency
5000–9999		1
10 000–14 999		5
15 000–19 999		9
20 000–24 999		3
25 000–29 999		2
30 000–34 999		2

a Copy and complete the class centre column for the frequency table.
b Show the information in a frequency histogram and polygon.

7 The frequency table below shows the marks achieved by Year 11 students on their English exam.

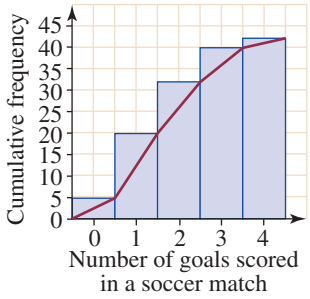
Class	Class centre	Frequency	Cumulative frequency
30–39		3	
40–49		6	
50–59		12	
60–69		15	
70–79		18	
80–79		10	

a Copy and complete the frequency table.
b Show the information on a cumulative frequency histogram and polygon.

8 Find the range of each of the following sets of scores.
a 28 24 26 24 25 29 22 27 25
b 118 2 56 45 72 43 69 84 159 0
c 1.9 0.7 0.5 0.8 1.1 1.5 1.4

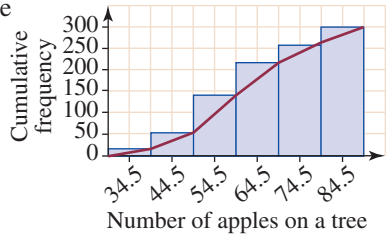
9 For each of the data sets in question **8** calculate:
i the median **ii** the lower quartile
iii the upper quartile **iv** the interquartile range.

10 The cumulative frequency histogram and polygon at right shows the number of goals scored by a soccer team in each match over a season.
Use the graph to calculate:



a the median
b the lower and upper quartiles
c the interquartile range.

11 The cumulative frequency histogram and polygon at right shows the number of apples on each tree in an orchard.
Use the graph to find:



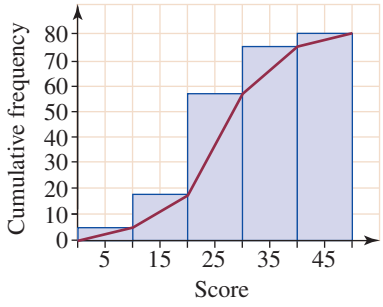
a the median
b the lower and upper quartiles
c the interquartile range.

12 Display the following scores in a stem-and-leaf plot.
45 21 38 46 42 41 42 49 35 29 24 28
36 21 38 45 44 40 29 28 35 35 33 38
40 41 48 39 34 38 45 28 23 29 30 40

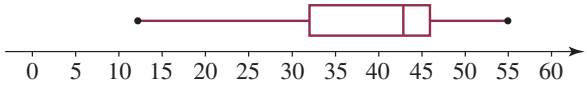
13 Use the stem-and-leaf plot drawn in the previous question to find:
a the range
b the median
c the interquartile range.

14 For the data set below, give a five-number summary.
24 53 91 57 29 69 29 15 84 6

15 Use the ogive to develop a five-number summary for the data set below.



16 For the box-and-whisker plot drawn at below:
a state the median
b calculate the range
c calculate the interquartile range.



17 The number of babies born each day at a hospital over a year is tabulated and the five-number summary is given below.
Lower extreme = 1 Upper quartile = 16
Lower quartile = 8 Upper extreme = 18
Median = 14
Show this information in a box-and-whisker plot.

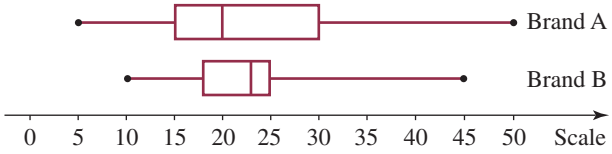
EXTENDED RESPONSE

1 The data set below shows the number of admissions to a hospital each day over a month.

25	20	33	41	15	18	24	40	12	29
30	38	26	20	17	23	10	11	16	23
22	27	14	11	12	14	32	24	29	33

a Using classes 10–14, 15–19, 20–24, etc., show this information in a frequency table.
b Draw a frequency histogram and polygon for the data.
c Draw a cumulative frequency histogram and polygon.
d State the range of the data set.
e Use the ogive to find:
  i the median **ii** the interquartile range.

2 The box-and-whisker plots below show the sales of two different brands of washing powder at a supermarket each day.



a State the range for Brand A.
b State the interquartile range for Brand A.
c State the range for Brand B.
d State the interquartile range for Brand B.
e Describe the spread of the sales for each brand of washing powder.

eBookplus
Digital doc
Test Yourself
doc-1580
Chapter 9

eBookplus ACTIVITIES

Are you ready?

Digital docs *(page 274)*
- SkillSHEET 9.3 (doc-1561): Reading a line graph
- SkillSHEET 9.4 (doc-1562): Presenting data as a dot plot
- SkillSHEET 9.5 (doc-1567): Producing a frequency table from a frequency histogram
- SkillSHEET 9.6 (doc-1571): Finding the median of a small data set
- SkillSHEET 9.7 (doc-1576): Presenting data as a stem-and-leaf plot

9A Frequency tables

Digital docs
- SkillSHEET 9.1 (doc-1558): Presenting data in a frequency table *(page 277)*
- Spreadsheet (doc-1559): Frequency tables *(page 277)*

9B Types of graphs

Tutorials
- WE4 int-2323: Display data on a sector graph. *(page 280)*
- WE7 int-2324: Display data on a radar chart. *(page 282)*

Digital docs
- SkillSHEET 9.2 (doc-1560): Reading a column graph *(page 283)*
- SkillSHEET 9.3 (doc-1561): Reading a line graph *(page 283)*
- SkillSHEET 9.4 (doc-1562): Presenting data as a dot plot *(page 283)*
- Spreadsheet (doc-1563): Column graphs (*page 283* and *page 285)*
- Spreadsheet (doc-1564): Line graphs *(page 283 and page 285)*
- Spreadsheet (doc-1565): Bar graphs *(page 283* and *page 286)*
- Spreadsheet (doc-1566): Sector graphs *(page 284 and page 286)*

9C Statistical graphs

Digital docs
- SkillSHEET 9.5 (doc-1567): Producing a frequency table from a frequency histogram *(page 289)*
- Spreadsheet (doc-1568): Histogram, frequency polygon *(page 289)*
- Spreadsheet (doc-1569): Histogram, frequency polygon (DIY) *(page 289)*
- WorkSHEET 9.1 (doc-1570): Apply knowledge of types of graphs to questions. *(page 292)*

9D Range and interquartile range

Tutorial
- WE15 int-2325: Learn to read the median and interquartile range from a cumulative frequency curve. *(page 298)*

Digital docs
- SkillSHEET 9.6 (doc-1571): Finding the median of a small data set *(page 299)*
- GC program — Casio (doc-1572): UV stats *(page 299)*
- GC program — TI (doc-1573): UV stats *(page 299)*
- Spreadsheet (doc-1574): Median *(page 301)*
- Spreadsheet (doc-1575): Median (DIY) *(page 301)*

9E Stem-and-leaf plots

Tutorial
- WE17 int-2326: Learn to draw a stem-and-leaf plot. *(page 307)*

Digital docs
- SkillSHEET 9.7 (doc-1576): Presenting data as a stem-and-leaf plot *(page 309)*

9F Five-number summaries

Interactivity
- int-0802: Boxplot and five-number summary *(page 313)*
- int-2401: Box-and-whisher plot *(page 313)*

Digital docs
- GC program — Casio (doc-1572): UV stats *(page 315)*
- GC program — TI (doc-1573): UV stats *(page 315)*
- WorkSHEET 9.2 (doc-1579): Apply knowledge of statistical graphs to problems. *(page 317)*

Chapter review

- Test yourself Chapter 9 (doc-1580): Take the end-of-chapter test to test your progress. *(page 321)*

To access eBookPLUS activities, log on to

www.jacplus.com.au